Convergence Rates for Exponentially Ill-Posed Inverse Problems with Impulsive Noise

Claudia König[†], Frank Werner[‡] and Thorsten Hohage[§]

November 30, 2015

This paper is concerned with exponentially ill-posed operator equations with additive impulsive noise on the right hand side, i.e. the noise is large on a small part of the domain and small or zero outside. It is well known that Tikhonov regularization with an L^1 data fidelity term outperforms Tikhonov regularization with an L^2 fidelity term in this case. This effect has recently been explained and quantified for the case of finitely smoothing operators. Here we extend this analysis to the case of infinitely smoothing forward operators under standard Sobolev smoothness assumptions on the solution, i.e. exponentially ill-posed inverse problems. It turns out that high order polynomial rates of convergence in the size of the support of large noise can be achieved rather than the poor logarithmic convergence rates typical for exponentially ill-posed problems. The main tools of our analysis are Banach spaces of analytic functions and interpolation-type inequalities for such spaces. We discuss two examples, the (periodic) backwards heat equation and an inverse problem in gradiometry.

Keywords: variational regularization, impulsive noise, spaces of analytic functions

AMS classification numbers: 65J20, 65K10, 65J22, 46B70

1 Introduction

In this work we analyze Tikhonov-type regularization for exponentially ill-posed problems where the data are corrupted by impulsive noise. We suppose that the measurements are described by functions $g^{\mathrm{obs}} = g^\dagger + \xi \in L^1(\mathbb{M})$ on a submanifold $\mathbb{M} \subset \mathbb{R}^d$ where $g^\dagger$ denotes the exact data and ξ the noise function. The noise ξ is called impulsive if $|\xi|$ is large on a small part of $\mathbb{M}$ and small or even zero elsewhere. Impulsive noise naturally occurs in digital image acquisition due to faulty memory locations or any kind of physical measurements with malfunctioning receivers. Inverse problems with impulsive noise have been studied extensively in the literature (see e.g. [2–4] and the

[†]Institute for Mathematical Stochastics, University of Göttingen, Germany and Inverse Problems in Biophysics Group, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany (claudia-juliane.koenig@mathematik.uni-goettingen.de)

[‡]Felix Bernstein Institute for Mathematical Statistics in the Biosciences, University of Göttingen, Germany and Inverse Problems in Biophysics Group, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany (Frank.Werner@mpibpc.mpg.de)

[§]Institute for Numerical and Applied Mathematics, University of Göttingen, Germany (hohage@math.uni-goettingen.de)

references therein), and several authors [1,3,5,7,10,26] also analyzed Tikhonov-type regularization where reconstructions are defined as minimizers of a generalized Tikhonov functional

$$\widehat{f}_\alpha \in \operatorname*{argmin}_{f \in D(F)} \left[\frac{1}{\alpha r} \left\| F(f) - g^{\mathrm{obs}} \right\|_{L^r(\mathbb{M})}^r + \mathcal{R}(f) \right]. \tag{1}$$

Here $\mathcal{X}$ is a Banach space, $F : D(F) \subset \mathcal{X} \to L^1(\mathbb{M})$ the forward operator, $\mathcal{R} : \mathcal{X} \to (-\infty, \infty]$ a convex penalty term stabilizing the reconstructions, $r \geq 1$, and $\alpha > 0$ is a regularization parameter. Most common examples for $\mathcal{R}$ are powers of Banach space norms $\mathcal{R}(f) = \frac{1}{q} \| f - f_0 \|_{\mathcal{X}}^q$ with $f_0 \in \mathcal{X}$ and $q \geq 1$ or total variation-type functionals.

It is well known from several numerical experiments [5,14,15] that the choice $r = 1$ in (1) leads to much better reconstructions than $r = 2$, and several efficient algorithms have been developed to minimize (1) with $r = 1$ (see [16,18,20,21,27,28]). The reconstruction improvements are even more striking if the forward problem arises from parameter identification problems in PDEs with smooth solutions [3].

The aforementioned analysis of (1) always requires $\| \xi \|_{L^r(\mathbb{M})} \to 0$, and is not able to fully explain the remarkable difference between $r = 1$ and $r = 2$. Recently, a new model to describe the impulsiveness of ξ has been proposed by the last two authors in [13]. Following this approach, we will assume that

$$\exists\, \mathbb{P} \in \mathfrak{B}(\mathbb{M}) : \qquad \| \xi \|_{L^1(\mathbb{M} \setminus \mathbb{P})} \leq \varepsilon, \qquad |\mathbb{P}| \leq \eta. \tag{2}$$

Here $\mathfrak{B}(\mathbb{M})$ denotes the Borel σ-algebra of $\mathbb{M}$ and $\varepsilon, \eta \geq 0$ are noise parameters. Under this model, the main result of [13] can be described as follows: If F maps $\mathcal{X}$ Lipschitz continuously into a Sobolev space $W^{k,p}(\mathbb{M})$ with $k > d/p$ and if the smoothness of the exact solution $f^\dagger$ is described by a variational inequality with index function φ (cf. (4) for details; this is ensured under standard conditions as we will discuss after Assumption 1 and in Corollary 2.3), then the Bregman distance between $f^\dagger$ and $\widehat{f}_\alpha$ is bounded by

$$\mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right) \leq C_1 \frac{\varepsilon}{\alpha} + C_2 \frac{\eta^{2\left(\frac{k}{d}+1\right) - \frac{2}{p}}}{\alpha^2} + C_3(-\varphi)^*\left(-\frac{1}{\alpha}\right) \tag{3}$$

with the Fenchel conjugate $(-\varphi)^*(s) := \sup_{t \geq 0}(st + \varphi(t))$ of $-\varphi$ (extended by ∞ on $(-\infty, 0)$). If we set $\eta = 0$ and hence $\varepsilon = \| \xi \|_{L^1(\mathbb{M})}$, this result basically coincides with the estimates proven in the literature on regularization with Banach norms [1,3,5,7,10]. Nevertheless, for mostly impulsive noise we expect $\varepsilon \approx 0$ and very small $\eta > 0$. In this situation the remarkable improvement is given by the higher exponent $2\left(\frac{k}{d} + 1\right) - \frac{2}{p}$ of η, which is fully determined by the smoothing properties of F.

In this work we study (1) in case that the range of F consists of analytic functions. This situation typically arises e.g. in parameter identification problems in PDEs with remote measurements. Thus we will obtain (3) for arbitrary k and p (probably influencing the constants), implying super-algebraic decay in η. Intuitively, as the smoothing property of the forward operator is of exponential type (if measured by the decay rate of the singular values), we may hope for an even better expression in η, namely an exponential one. Taking into account that in exponentially ill-posed cases the function φ is typically only logarithmic, this then implies logarithmic rates in ε, but still polynomial rates in η under an optimal choice of α.

This study is organized as follows: In the following § 2 we recall and adapt general convergence rate results for Tikhonov regularization (1) under variational source conditions to our noise model in combination with a new interpolation-type inequality. The variational source condition will be validated under classical spectral source conditions. § 3 contains the construction and properties of Banach spaces $\mathcal{A}^\lambda(\mathbb{M})$ of analytic functions on three manifolds $\mathbb{M}$ (circles, intervals, and spheres) as well as certain interpolation-type inequalities for these spaces. The index λ is always a weight function characterizing the growth of analytic extensions. If the forward operator additionally maps Lipschitz continuously into $\mathcal{A}^\lambda(\mathbb{M})$, then the general convergence analysis from § 2 is applicable. This mapping property is then verified for two practical examples in § 4, the (periodic) backwards heat equation and an inverse problem in gradiometry. We end this paper with some conclusions in § 5.

2 Generalized Tikhonov regularization

This section is devoted to the analysis of Tikhonov regularization (1). For the whole section let $\mathcal{X}$ be a Banach space, $r \geq 1$ and $F : \mathcal{D}(F) \subset \mathcal{X} \to L^r(\mathbb{M})$ an operator.

As discussed in [13] a minimizer $\widehat{f_\alpha}$ of (1) exists under reasonable assumptions on F and $\mathcal{R}$. Furthermore, it can be proven by standard arguments that $\widehat{f_\alpha}$ is stable w.r.t. g^{obs} in a suitable sense and converges to $f^\dagger$ as $\|\xi\|_{L^r(\mathbb{M})} \to 0$ if $\alpha = \alpha\left(\|\xi\|_{L^r(\mathbb{M})}\right)$ is chosen such that $\|\xi\|^r_{L^r(\mathbb{M})}/\alpha \to 0$ and $\alpha \to 0$. For details we refer to [13, Sec. 2.1] and the references therein.

2.1 Assumptions and merit discussion

In the following we state and discuss our assumptions used to prove rates of convergence. As usual an variational regularization theory will establish convergence estimates w.r.t. the Bregman distance $\mathcal{D}\left(\widehat{f_\alpha}, f^\dagger\right)$ defined by

$$\mathcal{D}\left(\widehat{f_\alpha}, f^\dagger\right) := \mathcal{R}(f) - \mathcal{R}\left(f^\dagger\right) - \langle f^*, f - f^\dagger \rangle,$$

where $f^* \in \partial\mathcal{R}(f^\dagger)$ is a subgradient of $\mathcal{R}$ at $f^\dagger$. Note that $\mathcal{D}\left(f, f^\dagger\right) = \left\|f - f^\dagger\right\|^2_{\mathcal{X}}$ if $\mathcal{R}(f) = \|f - f_0\|^2_{\mathcal{X}}$ and $\mathcal{X}$ is a Hilbert space.

Moreover, as in many other recent papers we will use an abstract smoothness assumption in the form of a variational inequality:

Assumption 1 (Variational source condition). *We assume that there exists $\beta > 0$ and a concave, increasing function $\varphi : [0, \infty) \to [0, \infty)$ with $\varphi(0) = 0$ such that*

$$\beta\mathcal{D}\left(f, f^\dagger\right) \leq \mathcal{R}(f) - \mathcal{R}(f^\dagger) + \varphi\left(\|F(f) - F(f^\dagger)\|^r_{L^r(\mathbb{M})}\right) \qquad for\ all\ f \in \mathcal{D}(F). \tag{4}$$

If $F : \mathcal{X} \to L^2(\mathbb{M})$ is a bounded linear operator between Hilbert spaces, $\mathcal{R}(f) = \|f - f_0\|^2_{\mathcal{X}}$, $f_0 \in \mathcal{X}$ and $r = 2$, then it has been shown [6, 8] that (4) is both sufficient and necessary for convergence rates of order $\varphi\left(\|\xi\|^2_{L^2(\mathbb{M})}\right)$. Note that this furthermore shows that a spectral source condition must imply a variational source condition with a function φ yielding the same rates of convergence in this situation. Another proof of this implication has been given by Flemming [6] for general φ.

We are mainly interested in exponentially ill-posed problems, where the most relevant case for the function φ in Assumption 1 is $\varphi(t) = C\varphi_p(t)$ with

$$\varphi_p(t) = (-\ln(t))^{-p}, \qquad t \leq \exp(-1). \tag{5}$$

This function can be defined on all of $\mathbb{R}$ by concave extension, but for our asymptotic studies only the behavior close to $t = 0$ is of relevance. As shown by Flemming [6], if $\mathcal{X}$ is a Hilbert space, $r = 2$ and F is bounded and linear, the spectral source condition $f^\dagger = \varphi_p(F^*F)w$ implies a variational source condition (4) with $\varphi = \beta'\varphi_{2p}$ and $\beta' > 0$ depending on $\|w\|$, $\|F\|$, and p.

For nonlinear operators F, the condition (4) can be seen as a combination of source and nonlinearity condition (see e.g. [22, Props. 3.35 and 3.38]). For inverse medium scattering problems it was recently shown with the help of complex geometrical optics solutions (cf. [12]) that a variational source condition with φ of the form (5) holds true if the solution belongs to a Sobolev ball, and p is an explicit function of the Sobolev index.

Both techniques for verifying Assumption 1 mentioned above yield (4) with $r = 2$. To obtain L^1-variational source conditions from L^2-variational source conditions, we will use the following smoothing properties of F, which also play a central role in the rest of our analysis:

Assumption 2 (smoothing properties). *Let $\mathcal{X}$ be a Banach space, $F : \mathcal{D}(F) \subset \mathcal{X} \to L^1(\mathbb{M})$ and $\mathcal{R} : \mathcal{X} \to (-\infty, \infty]$ a convex, lower semi-continuous penalty functional. Assume that there are $\delta_0 > 0$, $q > 1$ and an increasing function $\gamma : [0, \delta_0] \to [0, \infty)$ with $\lim_{\delta \to 0} \gamma(\delta) = 0$ such that*

$$\|F(f) - g^\dagger\|_{L^\infty(\mathbb{M})} \leq \gamma(\delta) \mathcal{D}\left(f, f^\dagger\right)^{\frac{1}{q}} + \frac{1}{\delta} \|F(f) - g^\dagger\|_{L^1(\mathbb{M})}. \tag{6}$$

for all $0 < \delta \leq \delta_0$ and all $f \in \mathcal{X}$.

Remark 2.1. *Suppose there exists a Banach space $\mathcal{Z} \subset L^\infty(\mathbb{M}) \cap L^1(\mathbb{M})$ such that*

(A) F maps Lipschitz continuously into $\mathcal{Z}$, i.e.

$$\|F(f) - F(f^\dagger)\|_{\mathcal{Z}} \leq C \mathcal{D}\left(f, f^\dagger\right)^{\frac{1}{q}} \tag{7}$$

with $C > 0$, $q > 1$ and

(B) there exists $\delta_0 > 0$ and an increasing function $\gamma : [0, \delta_0) \to [0, \infty)$ with $\lim_{\delta \to 0} \gamma(\delta) = 0$ such that for all $0 < \delta \leq \delta_0$ and all $g \in \mathcal{Z}$

$$\|g\|_{L^\infty(\mathbb{M})} \leq \gamma(\delta) \|g\|_{\mathcal{Z}} + \frac{1}{\delta} \|g\|_{L^1(\mathbb{M})}. \tag{8}$$

Inserting (7) in (8) with $g = F(f) - g^\dagger$ proves that Assumption 2 holds true with $\gamma(\delta)$ in (6) replaced by $C\gamma(\delta)$.

An inequality of the form (8) is shown in [13, Lem. 3.2] in the case of Sobolev spaces on a bounded Lipschitz domain $\mathbb{M} \subset \mathbb{R}^d$, $\mathcal{Z} = W^{k,p}(\mathbb{M})$ with $k > \frac{d}{p}$. In this case we obtained $\gamma(\delta) = c_1 \delta^{\frac{k}{d} - \frac{1}{p}}$ where $c_1 > 0$ is a constant depending only on $\mathbb{M}, k$ and p, but not on $\delta \in (0, \delta_0)$. As (8) holds for any $0 < \delta \leq \delta_0$, it can (in this special case of γ) be reformulated as

$$\|g\|_{L^\infty(\mathbb{M})} \leq \inf_{0 < \delta \leq \delta_0} \left[c_1 \delta^{\frac{k}{d} - \frac{1}{p}} \|g\|_{W^{k,p}(\mathbb{M})} + \frac{1}{\delta} \|g\|_{L^1(\mathbb{M})} \right] \leq C \|g\|_{W^{k,p}(\mathbb{M})}^{\frac{1}{u+1}} \|g\|_{L^1(\mathbb{M})}^{\frac{u}{u+1}} \tag{9}$$

with $u := k/d - 1/p$. If $\|g\|_{L^1(\mathbb{M})} \leq c_1 u \delta_0 \|g\|_{W^{k,p}(\mathbb{M})}$, this follows from calculating the infimum using differentiation w.r.t. δ. However, since $W^{k,p}(\mathbb{M})$ is continuously embedded into $L^1(\mathbb{M})$, (9) holds true for all g, with a larger constant C if $c_1 u \delta_0$ is smaller than the embedding constant. Consequently, (9) is of the form of a classical interpolation inequality and (8) can be interpreted as an interpolation inequality as well.

The more general formulations (6) and (8) allow for a wider class of spaces $\mathcal{Z}$, in particular spaces of analytic functions, which are more suitable for exponentially smoothing operators F. As in (9) the asymptotic behavior of $\gamma(\delta)$ as $\delta \searrow 0$ will depend on the smoothing properties of F, more precisely the more smoothing F the faster $\gamma(\delta) \searrow 0$ as $\delta \searrow 0$. In § 3 we will introduce spaces $\mathcal{Z}$ which guarantee (8) with exponentially decaying γ.

Now we are in position to prove the following proposition which validates Assumption 1 under a spectral source condition. As we are mainly interested in exponentially ill-posed problems, we restrict ourselves to logarithmic source conditions, which are appropriate in this case.

Proposition 2.2. *Let $\mathcal{X}$ be a Hilbert space, $\mathbb{M} \subset \mathbb{R}^d$ and choose $\mathcal{R}(f) = \|f\|_{\mathcal{X}}^2$, $r = 1$. Suppose that F satisfies Assumption 2 with $q = 2$ and (2) holds true.*

1. For all $\alpha > 0$, $0 \leq \eta \leq \delta_0/2$, and $\varepsilon \geq 0$ the estimator $\widehat{f}_\alpha$ from (1) is bounded by

$$\left\| \widehat{f}_\alpha \right\|_{\mathcal{X}} \leq \|f^\dagger\|_X + \sqrt{\frac{2\varepsilon}{\alpha}} + \frac{2\eta\gamma(2\eta)}{\alpha}. \tag{10}$$

2. Suppose $f^\dagger$ satisfies the variational source condition (Assumption 1) with $r = 2$, α is chosen such that

$$\sqrt{\frac{2\varepsilon}{\alpha}} + \frac{2\eta\gamma(2\eta)}{\alpha} \le C_{\mathrm{e}} \tag{11}$$

for all ε, η with some constant C_{e}, and F satisfies the Lipschitz condition (7). Then there exists $C_{\mathrm{v}} > 0$ independent of α, ε and η such that $f^\dagger$ satisfies the variational source condition (4) with $r = 1$ for all $f = \widehat{f}_\alpha$ and φ replaced by $\varphi(C_{\mathrm{v}}\cdot)$.

Proof. Part 1: Using the minimizing property

$$\frac{1}{\alpha}\left\|F(\widehat{f}_\alpha) - g^{\mathrm{obs}}\right\|_{L^1(\mathbb{M})} + \mathcal{R}(\widehat{f}_\alpha) \le \frac{1}{\alpha}\|\xi\|_{L^1(\mathbb{M})} + \mathcal{R}(f^\dagger) \tag{12}$$

and the noise model (2) we find the estimate

$$\left\|F(\widehat{f}_\alpha) - g^{\mathrm{obs}}\right\|_{L^1(\mathbb{M}\backslash\mathbb{P})} + \alpha\left\|\widehat{f}_\alpha\right\|_{\mathcal{X}}^2 \le \varepsilon + \alpha\left\|f^\dagger\right\|_{\mathcal{X}}^2 + \left\|F(\widehat{f}_\alpha) - g^\dagger\right\|_{L^1(\mathbb{P})} \tag{13}$$

Using (6) with $\delta = 2\eta$ and $\mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right)^{1/q} = \|\widehat{f}_\alpha - f^\dagger\| \le \|\widehat{f}_\alpha\| + \|f^\dagger\|$ (since $\mathcal{X}$ is a Hilbert space) it follows

$$\left\|F(\widehat{f}_\alpha) - g^\dagger\right\|_{L^1(\mathbb{P})} \le \frac{1}{2}\left\|F(\widehat{f}_\alpha) - g^\dagger\right\|_{L^1(\mathbb{M})} + \eta\gamma(2\eta)\left(\left\|\widehat{f}_\alpha\right\|_{\mathcal{X}} + \|f^\dagger\|_{\mathcal{X}}\right)$$

which then implies

$$\left\|F(\widehat{f}_\alpha) - g^\dagger\right\|_{L^1(\mathbb{P})} \le \left\|F(\widehat{f}_\alpha) - g^\dagger\right\|_{L^1(\mathbb{M}\backslash\mathbb{P})} + 2\eta\gamma(2\eta)\left(\left\|\widehat{f}_\alpha\right\|_{\mathcal{X}} + \|f^\dagger\|_{\mathcal{X}}\right). \tag{14}$$

As $\left|\left\|F(\widehat{f}_\alpha) - g^{\mathrm{obs}}\right\|_{L^1(\mathbb{M}\backslash\mathbb{P})} - \left\|F(\widehat{f}_\alpha) - g^\dagger\right\|_{L^1(\mathbb{M}\backslash\mathbb{P})}\right| \le \varepsilon$, we find from inserting (14) into (13) that

$$\left\|\widehat{f}_\alpha\right\|_{\mathcal{X}}^2 \le \frac{2\varepsilon}{\alpha} + \|f^\dagger\|_{\mathcal{X}}^2 + \frac{2\eta\gamma(2\eta)}{\alpha}\left(\left\|\widehat{f}_\alpha\right\|_{\mathcal{X}} + \|f^\dagger\|_{\mathcal{X}}\right).$$

Now we add $-\frac{2\eta\gamma(2\eta)}{\alpha}\left\|\widehat{f}_\alpha\right\|_{\mathcal{X}} + \frac{\eta^2\gamma(2\eta)^2}{\alpha^2}$ on both sides, complete the squares and obtain

$$\left(\left\|\widehat{f}_\alpha\right\| - \frac{\eta\gamma(2\eta)}{\alpha}\right)^2 \le \frac{2\varepsilon}{\alpha} + \left(\|f^\dagger\| + \frac{\eta\gamma(2\eta)}{\alpha}\right)^2 \le \left(\sqrt{\frac{2\varepsilon}{\alpha}} + \|f^\dagger\| + \frac{\eta\gamma(2\eta)}{\alpha}\right)^2.$$

Taking square roots yields the desired estimate.

Part 2: Note that $\left\|F(f) - g^\dagger\right\|_{L^2(\mathbb{M})}^2 \le \left\|F(f) - g^\dagger\right\|_{L^\infty(\mathbb{M})}\left\|F(f) - g^\dagger\right\|_{L^1(\mathbb{M})}$. Furthermore (6) together with (7) yields

$$\left\|F(f) - g^\dagger\right\|_{L^\infty(\mathbb{M})} \le \gamma(\delta)\left\|f - f^\dagger\right\|_{\mathcal{X}} + \frac{1}{\delta}\left\|F(f) - g^\dagger\right\|_{L^1(\mathbb{M})}$$

$$\le \left(\gamma(\delta) + \frac{C}{\delta}\right)\left\|f - f^\dagger\right\|_{\mathcal{X}}$$

for arbitrary δ. Together with the first part this shows that $\|F(f) - g^\dagger\|_{L^2(\mathbb{M})}^2 \le C_{\mathrm{v}}\|F(f) - g^\dagger\|_{L^1(\mathbb{M})}$ with $C_{\mathrm{v}} := (\gamma(\delta) + C/\delta)(2\|f^\dagger\|_{\mathcal{X}} + C_{\mathrm{e}})$, which yields the assertion. $\qquad\square$

As mentioned after (5), spectral logarithmic source conditions imply variational logarithmic source conditions. As $\varphi_{2p}(C_{\mathrm{v}}x) = \varphi_{2p}(x) + o(1)$ as $x \searrow 0$, we obtain:

Corollary 2.3. *If $\mathcal{X}$ is a Hilbert space, $F : \mathcal{X} \to L^2(\mathbb{M})$ is linear, the spectral source condition $f^\dagger = \varphi_p(F^*F)w$ is satisfied, and (11) holds true. Then the variational source condition (4) holds true for $f = \widehat{f}_\alpha$ with $r = 1$, and $\varphi = \beta'\varphi_{2p}$ where $\beta' > 0$ depends only on $\|w\|_{\mathcal{X}}$, $\|F\|$, p, δ_0, and C_{e}.*

2.2 Rates of convergence

Now we will present error estimates for the minimizers $\widehat{f}_\alpha$ in (1). Our analysis uses the observation by Grasmair [10] that the approximation error can be bounded by the Fenchel conjugate function of $-\varphi$. Let us introduce a function ψ by

$$\psi(\alpha) := (-\varphi)^* \left(-\frac{1}{\alpha}\right) = \sup_{\tau \geq 0} \left(\varphi(\tau) - \frac{\tau}{\alpha}\right), \qquad \alpha > 0. \tag{15}$$

Note from this definition that ψ is concave and non-decreasing.

For φ_p defined in (5) one obtains $\psi_p(\alpha) = \left(\ln \frac{1}{\alpha p}\right)^{-p} (1 + o(1))$ as $\alpha \to 0$.

We are now in a position to formulate and prove the main result of this section, which closely follows [13, Thm. 3.4]:

Theorem 2.4. *Let $\mathcal{X}$ be a Banach space, $F : \mathcal{D}(F) \subset \mathcal{X} \to L^1(\mathbb{M})$ be an operator and $\mathcal{R} : \mathcal{X} \to (-\infty, \infty]$ a convex, lower semi-continuous functional. Let $\widehat{f}_\alpha$ be a solution of (1) with $r = 1$ and suppose that the noise function ξ fulfills (2). If Assumptions 1 and 2 hold true with $q > 1$ (actually (4) is only needed for $f = \widehat{f}_\alpha$), then for any $b > 1$ there exists $\eta_0 > 0$ such that*

$$\beta \mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right) \leq \frac{2q'}{\alpha}\epsilon + \beta^{1-q'}\left(\frac{2}{\alpha}\eta\gamma\left(\frac{2b}{b-1}\eta\right)\right)^{q'} + q'\psi(b\alpha), \tag{16}$$

$$\frac{1}{b}\|F(\widehat{f}_\alpha) - g^\dagger\|_{L^1(\mathbb{M})} \leq 4q'\epsilon + 2\frac{\left(2\eta\gamma\left(\frac{2b}{b-1}\eta\right)\right)^{q'}}{(\beta\alpha)^{q'-1}} + 2\alpha q'\psi(2b\alpha) \tag{17}$$

for all $0 < \eta \leq \eta_0$, $\epsilon \geq 0$ and $\alpha > 0$. Here and below q' denotes the conjugate index of q, i.e. $\frac{1}{q} + \frac{1}{q'} = 1$.

Proof. By assumption there is a measurable $\mathbb{P}$ such that $\|\xi\|_{L^1(\mathbb{M}\setminus\mathbb{P})} \leq \epsilon$ and $|\mathbb{P}| \leq \eta$. Thus we have for any $g \in L^1(\mathbb{M})$ that

$$\|g - g^\dagger\|_{L^1(\mathbb{M})} + \|\xi\|_{L^1(\mathbb{M})} = \|g - g^\dagger\|_{L^1(\mathbb{M}\setminus\mathbb{P})} + \|g - g^\dagger\|_{L^1(\mathbb{P})} + \|\xi\|_{L^1(\mathbb{M})}$$

$$\leq \left(\|g - g^{\mathrm{obs}}\|_{L^1(\mathbb{M}\setminus\mathbb{P})} + \|\xi\|_{L^1(\mathbb{M}\setminus\mathbb{P})}\right) + \|g - g^\dagger\|_{L^1(\mathbb{P})} + \|\xi\|_{L^1(\mathbb{M})}$$

$$\leq \|g - g^{\mathrm{obs}}\|_{L^1(\mathbb{M})} + 2\|\xi\|_{L^1(\mathbb{M}\setminus\mathbb{P})} + 2\|g - g^\dagger\|_{L^1(\mathbb{P})}$$

$$\leq \|g - g^{\mathrm{obs}}\|_{L^1(\mathbb{M})} + 2\epsilon + 2|\mathbb{P}|\|g - g^\dagger\|_{L^\infty(\mathbb{M})}. \tag{18}$$

For $g = F(\widehat{f}_\alpha)$ we will now apply the interpolation inequality (8) with $\delta := \frac{2b}{b-1}\eta$. Given $b > 1$ we set $\eta_0 := \frac{b-1}{2b}\delta_0$ and obtain for any $0 < \eta \leq \eta_0$ that

$$\frac{1}{b}\|F(\widehat{f}_\alpha) - g^\dagger\|_{L^1(\mathbb{M})} + \|\xi\|_{L^1(\mathbb{M})}$$

$$\leq \|F(\widehat{f}_\alpha) - g^{\mathrm{obs}}\|_{L^1(\mathbb{M})} + 2\left(\epsilon + \eta\gamma\left(\frac{2b}{b-1}\eta\right)\mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right)^{\frac{1}{q}}\right). \tag{19}$$

Now Assumption 1 implies

$$
\beta \mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right)
$$

$$
\overset{(4)}{\leq} \mathcal{R}(\widehat{f}_\alpha) - \mathcal{R}(f^\dagger) + \varphi\left(\|F(\widehat{f}_\alpha) - F(f^\dagger)\|_{L^1(\mathbb{M})}\right)
$$

$$
\overset{(12)}{\leq} \frac{1}{\alpha}\|\xi\|_{L^1(\mathbb{M})} - \frac{1}{\alpha}\|F(\widehat{f}_\alpha) - g^{\mathrm{obs}}\|_{L^1(\mathbb{M})} + \varphi\left(\|F(\widehat{f}_\alpha) - F(f^\dagger)\|_{L^1(\mathbb{M})}\right)
$$

$$
\overset{(19)}{\leq} \frac{2}{\alpha}\left(\epsilon + \eta\gamma\left(\frac{2b}{b-1}\eta\right)\mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right)^{\frac{1}{q}}\right) - \frac{1}{b\alpha}\|F(\widehat{f}_\alpha) - g^\dagger\|_{L^1(\mathbb{M})}
$$
$$
+ \varphi\left(\|F(\widehat{f}_\alpha) - F(f^\dagger)\|_{L^1(\mathbb{M})}\right) \tag{20}
$$

$$
\leq \frac{2}{\alpha}\epsilon + \frac{2\eta}{\alpha}\gamma\left(\frac{2b}{b-1}\eta\right)\mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right)^{\frac{1}{q}} + \psi(b\alpha).
$$

The second term on the right-hand side can be handled by Young's inequality $xy \leq x^{q'}/q' + y^q/q$ with $x = \beta^{-1/q}\frac{2\eta}{\alpha}\gamma\left(\frac{2b}{b-1}\eta\right)$, which implies that (16) holds true.

Since $\mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right) \geq 0$, we derive from (20) and (16) in a similar fashion that

$$
\frac{1}{2b\alpha}\|F(\widehat{f}_\alpha) - g^\dagger\|_{L^1(\mathbb{M})}
$$

$$
\overset{(20)}{\leq} \frac{2}{\alpha}\left(\epsilon + \eta\gamma\left(\frac{2b}{b-1}\eta\right)\mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right)^{\frac{1}{q}}\right) - \frac{1}{2b\alpha}\|F(\widehat{f}_\alpha) - g^\dagger\|_{L^1(\mathbb{M})}
$$
$$
+ \varphi\left(\|F(\widehat{f}_\alpha) - F(f^\dagger)\|_{L^1(\mathbb{M})}\right)
$$

$$
\leq \frac{2}{\alpha}\epsilon + \frac{2}{\alpha}\eta\gamma\left(\frac{2b}{b-1}\eta\right)\mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right)^{\frac{1}{q}} + \psi(2b\alpha)
$$

$$
\leq \frac{2}{\alpha}\epsilon + \frac{1}{q'\beta^{q'-1}}\left(\frac{2}{\alpha}\eta\gamma\left(\frac{2b}{b-1}\eta\right)\right)^{q'} + \frac{1}{q}\beta\mathcal{D}\left(\widehat{f}_\alpha, f^\dagger\right) + \psi(2b\alpha)
$$

$$
\overset{(16)}{\leq} \frac{2q'}{\alpha}\epsilon + \beta^{1-q'}\left(\frac{2}{\alpha}\eta\gamma\left(\frac{2b}{b-1}\eta\right)\right)^{q'} + q'\psi(2b\alpha),
$$

where we used that ψ is monotonically increasing. This immediately proves (17). □

Remark 2.5. *The optimal α can be obtained by balancing out the terms, cf. [13, Thm. 2.3, 3.]. For simplicity we will do this only in the specific examples in § 4.*
Note that the phenomenon of exact penalization (cf. [1]) will occur if $\varphi(x) = c \cdot x$ with some $c > 0$, as then $\psi(\alpha) = 0$ if $\alpha \leq 1/c$ and $\psi(\alpha) = \infty$ otherwise. Consequently, in this situation the optimal parameter is $\alpha = 1/c$, and for noise-free data this already yields an exact reconstruction.

Remark 2.6. *All choices of α for which the right-hand side of (16) tends to 0 (e.g. a-priori choices which approximately minimize this right-hand side) satisfy the condition (11).*

3 Spaces of analytic functions and L^∞-L^1 interpolation inequalities

Let $\mathcal{A}(\mathbb{M})$ denote the set of all analytic functions on $\mathbb{M}$. In this section we will introduce spaces $\mathcal{A}^\lambda(\mathbb{M}) \subsetneq \mathcal{A}(\mathbb{M})$ of analytic functions on the manifolds $\mathbb{M} = \mathbb{R}/2\pi\mathbb{Z}, (-1,1)$, and SS^2 and establish interpolation inequalities for spaces $L^\infty(\mathbb{M}) \subset L^1(\mathbb{M}) \subset \mathcal{A}^\lambda(\mathbb{M})$.

General strategy

We will introduce classes of analytic functions $\mathcal{A}^\lambda(\mathbb{M})$ which are indexed by an increasing function $\lambda : [0, \infty) \to \mathbb{R} \cup \{\infty\}$. λ will be a bound on the growth rate of analytic extensions of functions $g \in \mathcal{A}^\lambda(\mathbb{M})$. Moreover, for each choice of $\mathbb{M}$ we will introduce a sequence of finite dimensional subspaces $\mathcal{U}_m \subset \mathcal{A}^\lambda(\mathbb{M})$, $m \in \mathbb{N}$ with corresponding orthogonal projections $P_m : L^2(\mathbb{M}) \to \mathcal{U}_m$ and prove exponential bounds of the approximation error of the form

$$\|g - P_m g\|_{L^\infty(\mathbb{M})} \le q_\lambda(m) \exp(-\lambda^*(m)) \|g\|_{\mathcal{A}^\lambda(\mathbb{M})} \tag{21}$$

for $m \ge m_0$ and $g \in \mathcal{A}^\lambda(\mathbb{M})$ with the Fenchel-conjugate $\lambda^*(s) := \sup_{x \ge 0}[rs - \lambda(r)]$ of the function λ (extended by ∞ on $(-\infty, 0)$) and a polynomial q_λ. As a second ingredient we will need that the projection operators P_m can be extended to $P_m : L^1(\mathbb{M}) \to L^\infty(\mathbb{M})$ and satisfy the operator norm bound

$$\|P_m g\|_{L^\infty(\mathbb{M})} \le \kappa_m \|g\|_{L^1(\mathbb{M})} \qquad \text{for all } g \in L^1(\mathbb{M}). \tag{22}$$

This is the case if P_m has an integral representation $(P_m g)(x) = \int_\mathbb{M} K_m(x, y) g(y) \, \mathrm{d}y$ and $\|K_m\|_{L^\infty(\mathbb{M} \times \mathbb{M})} = \kappa_m$. The kernel K_m is given by $K_m(x, y) = \sum_{j=1}^{N_m} g_j(x) g_j(y)$ for any orthonormal basis $\{g_1, \dots, g_{N_m}\} \subset \mathcal{U}_m$. With the help of (21) and (22) we obtain

$$\begin{aligned}\|g\|_{L^\infty(\mathbb{M})} &\le \|g - P_m g\|_{L^\infty(\mathbb{M})} + \|P_m g\|_{L^\infty(\mathbb{M})} \\ &\le q_\lambda(m) \exp(-\lambda^*(m)) \|g\|_{\mathcal{A}^\lambda(\mathbb{M})} + \kappa_m \|g\|_{L^1(\mathbb{M})}\end{aligned}$$

for $m \ge m_0$ and $g \in \mathcal{A}^\lambda(\mathbb{M})$. We then choose $m(\delta)$ (decreasing in $\delta > 0$) such that $\kappa_{m(\delta)} \le \delta^{-1}$. This yields the interpolation inequality

$$\begin{aligned}\|g\|_{L^\infty(\mathbb{M})} &\le \gamma(\delta) \|g\|_{\mathcal{A}^\lambda(\mathbb{M})} + \frac{1}{\delta} \|g\|_{L^1(\mathbb{M})} \quad \text{with} \\ \gamma(\delta) &:= q_\lambda(m(\delta)) \exp(-\lambda^*(m(\delta)))\end{aligned} \tag{23}$$

for all $g \in \mathcal{A}^\lambda(\mathbb{M})$ and $\delta \le \delta_0$ if $m(\delta_0) = m_0$ (see (8)).

3.1 Case $\mathbb{M} = \mathbb{R}/2\pi\mathbb{Z}$

Note that any 2π-periodic real-analytic function $g : \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{R}$ can be extended to a holomorphic function $\tilde{g}$ on some strip $S_B = \{z \in \mathbb{C} \mid |\Im(z)| < B\}$ with $B > 0$. We may measure the smoothness of g in terms of B and the growth of $|\tilde{g}(x + \mathrm{i}y)|$ as $|y| \nearrow B$. This leads to the following definition (cf. [17, Sec. 11] for the special case that λ is an indicator function):

Definition 3.1. *Let $\lambda : [0, \infty) \to \mathbb{R} \cup \{\infty\}$ be an increasing weight function with positive $B_\lambda := \sup\{r \in [0, \infty) \mid \lambda(r) < \infty\}$. With the holomorphic extension $\tilde{g}$ of g on B_λ introduced above we define the space $\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})$ by*

$$\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z}) := \left\{ g \in C(\mathbb{R}/2\pi\mathbb{Z}) \mid \tilde{g} \text{ exists on } S_{B_\lambda}, \quad \|g\|_{\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})} < \infty \right\},$$

$$\|g\|_{\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})} := \sup_{z \in S_{B_\lambda}} \left[\exp(-\lambda(|\Im(z)|)) |\tilde{g}(z)| \right].$$

Theorem 3.2. *Let $\lambda : [0, \infty) \to \mathbb{R} \cup \{\infty\}$ be non-decreasing with $B_\lambda > 0$.*

1. *The space $\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})$ equipped with $\|\cdot\|_{\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})}$ is a Banach space.*

2. *The Fourier coefficients $\hat{g}(n)$ of $g \in \mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})$ satisfy*

$$|\hat{g}(n)| \le \exp(-\lambda^*(|n|)) \|g\|_{\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})}, \qquad n \in \mathbb{Z}. \tag{24}$$

3. If P_m denotes the L^2-orthogonal projection onto span $\{\exp(\mathrm{i}n\cdot) \mid |n| \leq m\}$, then there exists $m_0 \in \mathbb{N}$ and a constant $c_\lambda > 0$ such that

$$\|(I - P_m)\,g\|_{L^\infty(-\pi,\pi)} \leq 2c_\lambda \exp\left(-\lambda^*(m)\right) \|g\|_{\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})} \tag{25}$$

for all $g \in \mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})$ and $m \geq m_0$.

4. There exists $\delta_0 > 0$ such that the interpolation inequality (23) holds true with $\mathbb{M} = \mathbb{R}/2\pi\mathbb{Z}$ and

$$\gamma(\delta) = c_\lambda \exp\left(-\lambda^*\left(\left\lfloor \frac{\pi}{\delta} - \frac{1}{2} \right\rfloor\right)\right)$$

for all $g \in \mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})$ and all $0 < \delta \leq \delta_0$. Here $\lfloor x \rfloor := \sup\{n \in \mathbb{Z} : n \leq x\}$ for $x \in \mathbb{R}$.

Proof. Part 1: Obviously, $\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})$ is a normed space. Since any Cauchy sequence in this spaces converges uniformly on compact subsets of B_λ, the pointwise limit is again holomorphic, so it belongs to $\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})$. This follows from the Cauchy integral formula and proves completeness.
Part 2: Let $g \in \mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})$ and denote by $\tilde{g}$ the holomorphic extension of g to S_{B_λ}. Since $\tilde{g}$ and $\tilde{g}(2\pi + \cdot)$ coincide on $\mathbb{R}$ by the periodicity of g, it follows from the identity principle that $\tilde{g}$ and $\tilde{g}(2\pi + \cdot)$ also coincide on S_{B_λ}, i.e. $\tilde{g}$ is 2π−periodic.

For $0 < R < B_\lambda$ consider the annulus $K_R := \{w \in \mathbb{C} \mid \exp(-R) < |w| < \exp(R)\}$. The map $\theta : S_R \to K_R, \theta(z) := \exp(\mathrm{i}z)$ is holomorphic for any $0 < R \leq B_\lambda$. The restriction of θ to the rectangle $D_R := \{z \in S_R \mid \Re(z) \in (-\pi, \pi]\}$ is bijective and $\theta(-\pi + \mathrm{i}y) = \theta(\pi + \mathrm{i}y) = -\exp(-y)$. Hence $h := \tilde{g} \circ \theta^{-1} : K_{B_\lambda} \to \mathbb{C}$ is continuous on the line segment $\{-\exp(-y) \mid -R < y < R\}$ and therefore holomorphic in the annulus K_{B_λ}. If follows that for all $w = \exp(\mathrm{i}x) \in SS^1, x \in (-\pi, \pi]$ we have

$$h(w) = g \circ \theta^{-1}(w) = g(x) = \sum_{n \in \mathbb{Z}} \hat{g}(n) \exp(\mathrm{i}nx) = \sum_{n \in \mathbb{Z}} \hat{g}(n) w^n. \tag{26}$$

Being holomorphic in the annulus K_{B_λ}, h has a Laurent series expansion which by the identity principle is uniquely determined by the series expansion (26) on SS^1, i.e. $h(w) = \sum_{n \in \mathbb{Z}} \hat{g}(n) w^n$ for all $w \in K_{B_\lambda}$. For any $s \in \mathbb{R}$ with $\exp(-B_\lambda) < s < \exp(B_\lambda)$ we have

$$|\hat{g}(n)| = \left| \frac{1}{2\pi i} \int_{|w|=s} h(w) \frac{\mathrm{d}w}{w^{n+1}} \right| \leq \max_{|w|=s} |h(w)| \cdot \frac{1}{s^n}.$$

But $\max_{|w|=s} |h(w)| = \max_{\Im(z)=\ln(1/s)} |\tilde{g}(z)|$, and hence with $r := \ln(1/s)$ this implies

$$|\hat{g}(n)| \leq \max_{\Im(z)=\pm r} |\tilde{g}(z)| \cdot \exp(-r|n|) \leq \|g\|_{\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})} \cdot \exp(-[r|n| - \lambda(r)])$$

for all $0 < r < B_\lambda$. Optimizing in r proves (24).
Part 3: The orthogonal projection is given by $P_m g(x) = \sum_{|n| \leq m} \hat{g}(n) \exp(\mathrm{i}nx)$, $x \in (-\pi, \pi)$. Using (24) we obtain

$$\|g - P_m g\|_{L^\infty(-\pi,\pi)} = \sup_{x \in (-\pi,\pi)} \left| \sum_{|n|>m} \hat{g}(n) \exp(\mathrm{i}nx) \right| \leq 2 \|g\|_{\mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})} \sum_{n=m+1}^\infty \exp(-\lambda^*(n))).$$

Recall that λ^* is always convex. As $\lambda(r) = \infty$ for $r < 0$, $\lambda^*(s) = \sup_{r \geq 0}[sr - \lambda(r)]$ is non-decreasing, and since $B_\lambda > 0$ it is easy to see that there exists $m_0 \in \mathbb{N}$ with $\lambda^*(m_0) > 0$ and $a := \lambda^*(m_0+1) - \lambda^*(m_0) > 0$. Thus the convexity of λ^* implies $a \leq \lambda^*(m+1) - \lambda^*(m)$ for all $m \geq m_0$. Consequently, for any $n > m \geq m_0$, we have $\lambda^*(n) - \lambda^*(m) = \sum_{j=m}^{n-1}(\lambda^*(j+1) - \lambda^*(j)) \geq (n-m)a$. Therefore,

$$\sum_{n=m+1}^\infty \exp(-[\lambda^*(n) - \lambda^*(m)]) \leq \sum_{n=m+1}^\infty \exp(-a(n-m)) = \frac{1}{\exp(a) - 1}$$

whenever $m \geq m_0$. This shows (25) with $c_\lambda := 2/(\exp(a) - 1)$.
Part 4: We follow our general strategy with $K_m(x,y) := \frac{1}{2\pi} \frac{\sin((m+1/2)(x-y))}{\sin((x-y)/2)}$ (Dirichlet kernel), $\kappa_m := \frac{2m+1}{2\pi}$, and $m(\delta) = \left\lfloor \frac{\pi}{\delta} - \frac{1}{2} \right\rfloor$. $\qquad\square$

3.2 Case $\mathbb{M} = (-1, 1)$

The main idea is to extend a real-analytic function g on $(-1, 1)$ to a holomorphic function $\tilde{g}$ on an ellipse with foci $\{-1, 1\}$ of the form

$$\mathcal{E}_r = \left\{ x + \mathrm{i}y \in \mathbb{C} \ \middle| \ x, y \in \mathbb{R}, \frac{x^2}{\cosh(r)^2} + \frac{y^2}{\sinh(r)^2} \le 1 \right\}, \qquad r > 0.$$

Definition 3.3. *Let* $\lambda : [0, \infty) \to \mathbb{R} \cup \{\infty\}$ *be an increasing weight function with positive* $B_\lambda := \sup \{ r \in [0, \infty) \mid \lambda(r) < \infty \}$. *We define the space* $\mathcal{A}^\lambda(-1, 1)$ *by*

$$\mathcal{A}^\lambda(-1, 1) := \left\{ g \in \mathcal{A}(-1, 1) \mid \tilde{g} \text{ exists on } \mathcal{E}_{B_\lambda}, \quad \|g\|_{\mathcal{A}^\lambda(-1,1)} < \infty \right\},$$

$$\|g\|_{\mathcal{A}^\lambda(-1,1)} := \sup_{0 < t < B_\lambda} \left[\exp(-\lambda(t)) \sup_{z \in \partial \mathcal{E}_t} |\tilde{g}(z)| \right].$$

Theorem 3.4. *Let* $\lambda : [0, \infty) \to \mathbb{R} \cup \{\infty\}$ *be non-decreasing with* $B_\lambda > 0$.

1. *The space* $\mathcal{A}^\lambda(-1, 1)$ *equipped with the norm* $\|\cdot\|_{\mathcal{A}^\lambda(-1,1)}$ *is a Banach space.*

2. *The coefficients* $a_n(g)$ *of* $g \in \mathcal{A}^\lambda(-1, 1)$ *with respect to the Chebychev polynomials satisfy*

$$|a_n(g)| \le 2 \exp(-\lambda^*(|n|)) \|g\|_{\mathcal{A}^\lambda(-1,1)}, \qquad n \in \mathbb{N}. \tag{27}$$

3. *If* P_m *denotes the* L^2*-orthogonal projection onto the space of polynomials of degree* $\le m$, *then there exists* $m_0 \in \mathbb{N}$ *and a constant* $c_\lambda > 0$ *such that*

$$\|(I - P_m)g\|_{L^\infty(-1,1)} \le c_\lambda \exp(-\lambda^*(m)) \|g\|_{\mathcal{A}^\lambda(-1,1)} \tag{28}$$

for all $g \in \mathcal{A}^\lambda(-1, 1)$ *and* $m \ge m_0$.

4. *There exists* $\delta_0 > 0$ *such that the interpolation inequality* (23) *holds true with* $\mathbb{M} = (-1, 1)$ *and*

$$\gamma(\delta) = c_\lambda \exp\left(-\lambda^* \left(\left\lfloor \sqrt{\frac{2}{\delta}} \right\rfloor - 1 \right) \right) \tag{29}$$

for all $g \in \mathcal{A}^\lambda(-1, 1)$ *and all* $\delta \le \delta_0$.

Proof. Part 1: This can be shown as in the proof of Theorem 3.2.

Part 2: The following approximation argument is taken from Kress [17, Thm. 11.7]. The function $\theta : K_{B_\lambda} \to \mathcal{E}_{B_\lambda}$, $\theta(w) := \frac{1}{2}\left(w + \frac{1}{w}\right)$ on an annulus K_{B_λ} as defined in the proof of Theorem 3.2 is surjective and holomorphic (but not injective).

Now let $g \in \mathcal{A}^\lambda(-1, 1)$ with holomorphic extension $\tilde{g} : \mathcal{E}_{B_\lambda} \to \mathbb{C}$. Define $h : K_{B_\lambda} \to \mathbb{C}$ by $h(w) := 2\tilde{g} \circ \theta(w)$. The holomorphic function h in the annulus K_{B_λ} can be expanded into a Laurent series $h(w) = \sum_{n \in \mathbb{Z}} a_n w^n$, $w \in K_{B_\lambda}$ with coefficients

$$a_n = \frac{1}{\pi i} \int_{|w| = \exp(r)} \tilde{g}\left(\frac{1}{2}\left(w + \frac{1}{w} \right) \right) \frac{\mathrm{d}w}{w^{n+1}}, \qquad -B_\lambda < r < B_\lambda, \ n \in \mathbb{Z}.$$

Substitution $\tilde{w} = \frac{1}{w}$ shows that $a_{-n} = a_n$. Therefore

$$h(w) = a_0 + \sum_{n \in \mathbb{N}} a_n \left(w^n + \frac{1}{w^n} \right), \qquad w \in K_{B_\lambda}.$$

For $|w| = 1$ we write $w = \exp(\mathrm{i}t)$ and obtain

$$\frac{1}{2}\left(w^n + \frac{1}{w^n} \right) = \cos(nt) = T_n(\cos(t)) = T_n\left(\frac{1}{2}\left(w + \frac{1}{w} \right) \right),$$

where T_n is the n−th Chebychev polynomial. For $z \in \mathcal{E}_{B_\lambda}$ we find $w \in K_{B_\lambda}$ with $z = \theta(w)$ to obtain

$$\tilde{g}(z) = \tilde{g} \circ \theta(w) = \frac{1}{2}h(w) = \frac{a_0}{2} + \sum_{n \in \mathbb{N}} a_n T_n(z).$$

The formula for a_n yields for any $n \in \mathbb{N}$ and $0 \le r < B_\lambda$ that

$$|a_n| \le \frac{2}{\exp(rn)} \max_{z \in \partial \mathcal{E}_r} |\tilde{g}(z)| \le 2 \exp(\lambda(r) - nr) \|g\|_{\mathcal{A}^\lambda(-1,1)} \tag{30}$$

where we used the definition of $\|\cdot\|_{\mathcal{A}^\lambda(-1,1)}$. Optimizing in r yields (27).

Part 3: As $P_m g(x) = a_0(g)/2 + \sum_{n=1}^m a_n(g)T_n(x)$ and $\|T_n\|_{L^\infty(-1,1)} = 1$, we obtain

$$\|g - P_m g\|_{L^\infty(-1,1)} = \sum_{n=m+1}^\infty |a_n(g)| \le 2 \sum_{n=m+1}^\infty \exp(-\lambda^*(n)) \|g\|_{\mathcal{A}^\lambda(-1,1)}. \tag{31}$$

The sum may be bounded as in the proof of Theorem 3.2.

Part 4: We again follow our general strategy. A complete orthonormal system of $\mathcal{U}_m = \text{span}\{x^0, \cdots, x^m\} \subset L^2(-1,1)$ is given by $\{\sqrt{j+1/2}\, p_j : j = 0, \ldots, m\}$ with the Legendre polynomials p_j. As $\|p_j\|_{L^\infty(-1,1)} = 1$, the supremum of the kernel $K(x,y) = \sum_{j=0}^m (j+1/2) p_j(x) p_j(y)$ is bounded by $\kappa_m = \sum_{j=0}^m (j+1/2) = (m+1)^2/2$. Hence, we choose $m(\delta) := \lfloor \sqrt{2/\delta} \rfloor - 1$. $\qquad\square$

3.3 Case $\mathbb{M} = SS^2$

Let $SS^2 := \{x \in \mathbb{R}^3 : |x|_2 = 1\}$ denote the unit sphere. Recall that if $\mathcal{P}_m$ denotes the space of polynomials in $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ of degree $\le m$, then the space of *spherical harmonics* $\mathcal{H}_m$ is defined by

$$\mathcal{H}_m := \left\{ p_{|SS^2} \big| p \in \mathcal{P}_m \text{ harmonic and } p \text{ is homogeneous of degree } m \right\}.$$

Here harmonic means $\Delta p = 0$ and homogeneous of degree m that $p(rx) = r^m p(x)$ for all $r > 0$ and $x \in \mathbb{R}^3$. We have the following decompositions as orthogonal direct sums with respect to $\langle \cdot, \cdot \rangle_{L^2(SS^2)}$ (see e.g. [24, Cpt. 4]):

$$L^2(SS^2) = \bigoplus_{l=0}^\infty \mathcal{H}_l, \qquad \mathcal{P}_m|_{SS^2} = \bigoplus_{l=0}^m \mathcal{H}_l \tag{32}$$

One has $\dim \mathcal{H}_m = 2m+1$ and $\dim \mathcal{P}_m|_{SS^2} = \sum_{l=0}^m (2l+1) = (m+1)^2$. The orthogonal projections of $L^2(SS^2)$ onto $\mathcal{H}_m$ and $\mathcal{P}_m|_{SS^2}$ will be denoted by Q_m and P_m, respectively. Spherical harmonics are closely related to Legendre polynomials p_m. Choose any orthonormal basis $(Y_j)_{j=1}^{2m+1}$ of $\mathcal{H}_m$ with respect to $\langle \cdot, \cdot \rangle_{L^2(SS^2)}$. The *addition formula* of the spherical harmonics (see e.g. [19, Thm. 2]) states that $\frac{2m+1}{4\pi} p_m(\langle x, y \rangle) = \sum_{j=1}^{2m+1} Y_j(x) Y_j(y)$, i.e.

$$(Q_m g)(x) = \frac{2m+1}{4\pi} \int_{SS^2} p_m(\langle x, y \rangle) g(y) \, \mathrm{d}y, \qquad x \in SS^2 \tag{33}$$

Let us introduce the averaging operator $M : C(SS^2) \to C(SS^2 \times [-1,1])$ by

$$(Mg)(x,t) := \begin{cases} \frac{1}{2\pi\sqrt{1-t^2}} \int_{\{y \in SS^2 | \langle y, x \rangle = t\}} g(y)\mathrm{d}y, & t \in (-1,1), \\ g(\pm x), & t = \pm 1. \end{cases}$$

Note that $(Mg)(x,t)$ is the average of f over a circle of radius $\sqrt{1-t^2}$ around x and that Mg is in fact continuous if f is continuous.

11

Definition 3.5. *Let $\lambda : [0, \infty) \to \mathbb{R} \cup \{\infty\}$ be an increasing weight function. We define the space $\mathcal{A}^\lambda(SS^2)$ by*

$$\mathcal{A}^\lambda(SS^2) := \left\{ g \in C(SS^2) \mid (Mg)(x, \cdot) \in \mathcal{A}^\lambda(-1, 1) \text{ for all } x \in SS^2, \|g\|_{\mathcal{A}^\lambda(SS^2)} < \infty \right\},$$

$$\|g\|_{\mathcal{A}^\lambda(SS^2)} := \sup_{x \in SS^2} \|Mg(x, \cdot)\|_{\mathcal{A}^\lambda(-1,1)}.$$

Theorem 3.6. *Let $\lambda : [0, \infty) \to \mathbb{R} \cup \{\infty\}$ be non-decreasing with $B_\lambda > 0$.*

1. *$\mathcal{A}^\lambda(SS^2)$ equipped with the norm $\| \cdot \|_{\mathcal{A}^\lambda(SS^2)}$ is a Banach space.*

2. *For all $g \in \mathcal{A}^\lambda(SS^2)$ we have*

$$\|Q_m g\|_{L^\infty(SS^2)} \leq \frac{2m+1}{4} \exp(-\lambda^*(m-1)) \|g\|_{\mathcal{A}^\lambda(SS^2)}. \tag{34}$$

3. *There exist constants $c_\lambda, d_\lambda > 0$ and $m_0 \in \{0, 1, \dots\}$ such that*

$$\|g - P_m g\|_{L^\infty(SS^2)} \leq (c_\lambda + m d_\lambda) \exp(-\lambda^*(m)) \|g\|_{\mathcal{A}^\lambda(SS^2)} \tag{35}$$

   for all $g \in \mathcal{A}^\lambda(SS^2)$ and $m \geq m_0$.

4. *There exists $\delta_0 > 0$ such that the interpolation inequality (23) holds true with $\mathbb{M} = SS^2$ and*

$$\gamma(\delta) = \left(c_\lambda + \sqrt{\frac{4\pi}{\delta}} d_\lambda \right) \exp\left(-\lambda^* \left(\left\lceil \sqrt{\frac{4\pi}{\delta}} \right\rceil - 1 \right) \right)$$

   for all $g \in \mathcal{A}^\lambda(SS^2)$ and all $0 < \delta \leq \delta_0$.

Proof. Part 1: Obviously $A^\lambda(SS^2)$ is a normed space. To show completeness, let (g_n) be a Cauchy-sequence in $A^\lambda(SS^2)$. Then (Mg_n) converges to some $G \in C(SS^2 \times [-1, 1])$ with $\sup_{x \in SS^2} \|G(x, \cdot)\|_{\mathcal{A}^\lambda(-1,1)} < \infty$ and (g_n) converges uniformly to $G(\cdot, 1)$. Since M is continuous, it follows that $G = MG(\cdot, 1)$.
Part 2: Due to (33) and the formula $\int_{SS^2} f(y) \, \mathrm{d}y = \int_{-1}^1 \frac{1}{\sqrt{1-t^2}} \int_{\{y \mid \langle y, x \rangle = t\}} f(y) \, \mathrm{d}y \, \mathrm{d}t$ we have

$$(Q_m g)(x) = \frac{2m+1}{2} \int_{-1}^1 p_m(t)(Mg)(x, t) \, \mathrm{d}t$$

$$= \frac{2m+1}{2} \int_0^\pi p_m(\cos(s))(Mg)(x, \cos(s)) \sin(s) \, \mathrm{d}s.$$

With the mapping $\theta(w) := \frac{1}{2}(w + \frac{1}{w})$ from the proof of Theorem 3.4 and the substitution $w = \exp(\mathrm{i}s)$ we obtain

$$(Q_m g)(x) = \frac{2m+1}{4} \Re \int_{\{w \in \mathbb{C} \mid |w| = 1\}} p_m(\theta(w)) (Mg)(x, \theta(w)) \, \mathrm{d}w.$$

We may deform the contour of integration $\{w \mid |w| = 1\}$ to any contour $\{w \mid |w| = \exp(r)\}$ with $|r| < B_\lambda$. To estimate $Q_m g$, we have to estimate the growth of the Legendre polynomials on $\mathcal{E}_r$. We use the identity $p_m(z) = \frac{1}{\pi} \int_0^\pi \left(z + \sqrt{z^2 - 1} \cos \varphi \right)^m \, \mathrm{d}\varphi$ (see [25, Cpt. 4]) to find that

$$p_m\left(\frac{1}{2} \left(w + \frac{1}{w} \right) \right) = \frac{1}{\pi} \int_0^\pi \left[w \frac{1 + \cos \varphi}{2} + \frac{1}{w} \frac{1 - \cos \varphi}{2} \right]^m \, \mathrm{d}\varphi$$

$$= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \left[w \cos^2 \psi + \frac{1}{w} \sin^2 \psi \right]^m \, \mathrm{d}\psi.$$

Therefore, $\sup_{z \in \partial \mathcal{E}_r} |p_m(z)| = \sup_{|w|=r} \left| p_m \left(\frac{1}{2} \left(w + \frac{1}{w} \right) \right) \right| \leq \exp(mr)$. It follows that

$$
\begin{aligned}
|(Q_m g)(x)| &\leq \frac{2m+1}{4} \sup_{0 \leq r < B_\lambda} \int_{|w|=\exp(r)} |p_m(\theta(w))| \, |(Mg)(x, \theta(w))| \, |\mathrm{d}w| \\
&\leq \frac{2m+1}{4} \sup_{r \geq 0} \left[\exp \left(r(m-1) - \lambda(r) \right) \right] \|g\|_{\mathcal{A}^\lambda(SS^2)} \\
&= \frac{2m+1}{4} \exp(-\lambda^*(m-1)) \|g\|_{\mathcal{A}^\lambda(SS^2)}.
\end{aligned}
$$

Part 3: As shown in the proof of Theorem 3.2 there exists $m_0 \in \{0, 1, \dots\}$ and $a > 0$ such that $\lambda^*(m_0) > 0$ and $\lambda^*(m+1) - \lambda^*(m) \geq a$ for all $m \geq m_0$. Using the identity $\sum_{j=0}^\infty r^j = 1/(1-r)$ for $|r| < 1$ and its derivative $\sum_{m=1}^\infty m r^{m-1} = (1-r)^{-2}$ we obtain

$$
\begin{aligned}
\|g - Pg\|_{L^\infty(SS^2)} &\leq \sum_{j=m+1}^\infty \|Q_m g\|_{L^\infty(SS^2)} \\
&\leq \sum_{j=m+1}^\infty \frac{2j+1}{4} \exp(-\lambda^*(j-1)) \|g\|_{\mathcal{A}^\lambda(SS^2)} \\
&\leq \exp(-\lambda^*(m)) \sum_{j=m+1}^\infty \frac{2j+1}{4} \exp(-\lambda^*(j-1) + \lambda^*(m)) \|g\|_{\mathcal{A}^\lambda(SS^2)} \\
&\leq \exp(-\lambda^*(m)) \|g\|_{\mathcal{A}^\lambda(SS^2)} \sum_{l=0}^\infty \left(\frac{l+1}{2} + \frac{2m+1}{4} \right) \exp(-al) \\
&= \exp(-\lambda^*(m)) \|g\|_{\mathcal{A}^\lambda(SS^2)} \left(\frac{1}{2(1-\exp(-a))^2} + \frac{2m+1}{4 - 4\exp(-a)} \right)
\end{aligned}
$$

This shows (35).

Part 4: By (32) and (33) the kernel of the L^2-orthogonal projection onto $\mathcal{P}_m|_{SS^2}$ is given by $K_m(x,y) = \frac{1}{4\pi} \sum_{l=0}^m (2l+1) p_l(\langle x, y \rangle)$. As $\|p_l\|_{L^\infty(-1,1)} = 1$, the supremum norm of K_m is bounded by $\kappa_m = \frac{1}{4\pi} \sum_{l=0}^m (2l+1) = \frac{1}{4\pi}(m+1)^2$. Therefore, we choose $m(\delta) = \lfloor \sqrt{4\pi/\delta} \rfloor - 1$. $\qquad \square$

4 Examples

In this section we show how our general techniques can be applied to practical examples. As we have seen in § 2, the key ingredients are the variational source condition (4) and the smoothness assumption (6). As the variational source condition can be validated using Proposition 2.2, we will focus on verifying (6). This is done exemplary for the forward operators connected to the periodic backwards heat equation and an inverse problem in satellite gradiometry on SS^2. Similar techniques should apply to many other exponentially ill-posed problems, e.g. inverse scattering problems.

4.1 Backwards heat equation

Let $f \in L^2(-\pi, \pi)$. Let $u : (-\pi, \pi) \times [0, \infty) \to \mathbb{R}$ be a solution of the periodic heat equation

$$
\begin{cases}
\frac{\partial u}{\partial t}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t) & \text{if } (x,t) \in (-\pi, \pi) \times (0, \infty) \\
u(x,0) = f(x) & \text{if } x \in (-\pi, \pi) \quad \text{(initial condition)} \\
u(-\pi, t) = u(\pi, t) & \text{if } t \in (0, \infty) \quad \text{(boundary condition)}
\end{cases}
\tag{36}
$$

It describes heat propagation on a circle parameterized by x as angular variable. The initial boundary value problem (36) has a unique solution given by

$$
u(x,t) = \sum_{n \in \mathbb{Z}} \exp \left(-n^2 t \right) \, \hat{f}(n) \exp \left(\mathrm{i} n x \right), \qquad (x,t) \in [-\pi, \pi] \times [0, \infty)
\tag{37}
$$

where $\hat{f}(n)$ are the Fourier coefficients of f.

Fix a time $\bar{t} > 0$. In the following we will study the forward operator F associated to the backwards heat equation, which consists in determining the initial heat distribution f from the measurement of the heat distribution $u(x, \bar{t})$ at time $\bar{t}$. This problem is known to be exponentially ill-posed, which can readily be seen from the decay of the singular values $\exp\left(-n^2 t\right)$ in (37).

Let us therefore define $T : L^2(-\pi, \pi) \to L^1(-\pi, \pi)$ by

$$(Tf)(x) := \sum_{n \in \mathbb{Z}} \exp\left(-n^2 \bar{t}\right) \ \hat{f}(n) \exp\left(inx\right), \qquad x \in [-\pi, \pi] \tag{38}$$

and set $g := u(\cdot, \bar{t})$.

Lemma 4.1. *Let T as in (38), set $\mathcal{X} := L^2(-\pi, \pi)$ and $\mathcal{R}(f) = \|f\|_{L^2(-\pi,\pi)}^2$.*

1. *Then $T : \mathcal{X} \to \mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})$ is bounded with*

$$\lambda(r) = \frac{r^2}{4\bar{t}}, \qquad r \in [0, \infty).$$

   Furthermore Assumption 2 holds true with $q = 2$ and

$$\gamma(\delta) \leq \frac{8 \max\left\{1, \bar{t}^{-1/2}\right\}}{\exp(3\bar{t}) - 1} \exp\left(-\left\lfloor \frac{\pi}{\delta} - \frac{1}{2} \right\rfloor^2 \bar{t}\right).$$

2. *If $f^\dagger \in H^p(-\pi, \pi)$, then Assumption 1 holds true with any bounded $D(F)$, $\mathbb{M} = (-\pi, \pi)$, $r = 1$, some $\beta, \beta' > 0$ and*
$$\varphi(\tau) = \beta' \varphi_p(\tau), \qquad \tau > 0.$$

Proof. Part 1: First we estimate $\|T\|_{\mathcal{X} \to \mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})}$. For $x \in \mathbb{R}, t > 0$ we will use the identity

$$\frac{1}{\sqrt{\pi}} \sum_{n \in \mathbb{Z}} \exp\left(-(n+x)^2 t\right) \leq \frac{1}{\sqrt{\pi}} \sum_{n \in \mathbb{Z}} \exp\left(-n^2 t\right) \leq \sqrt{2} \max\left\{1, \frac{1}{\sqrt{t}}\right\}, \tag{39}$$

which follows from Poisson's summation formula (cf. [23, Cpt. 5, Thm. 3.1]).

Applying Cauchy-Schwarz, $\sup_{\Im(z)=\pm r} |\exp(inz)| \leq \exp(|n|r)$ and (39) with $x = \pm\frac{r}{2\bar{t}}$ we find

$$\sup_{\Im(z)=\pm r} |Tf(z)| \leq \sum_{n \in \mathbb{Z}} \exp(-n^2 \bar{t}) \exp(|n|r) \frac{1}{\sqrt{2\pi}} \|f\|_{\mathcal{X}}$$

$$= \frac{1}{\sqrt{2\pi}} \exp(\lambda(r)) \|f\|_{\mathcal{X}} \sum_{n \in \mathbb{Z}} \exp\left(-\left(|n| - \frac{r}{2\bar{t}}\right)^2 \bar{t}\right)$$

The sum on the right hand side can be bounded as follows:

$$\sum_{n \in \mathbb{Z}} \exp\left(-\left(|n| - \frac{r}{2\bar{t}}\right)^2 \bar{t}\right) \leq \sum_{n \in \mathbb{Z}} \left[\exp\left(-\left(n - \frac{r}{2\bar{t}}\right)^2 \bar{t}\right) + \exp\left(-\left(n + \frac{r}{2\bar{t}}\right)^2 \bar{t}\right)\right]$$

$$\leq 2 \sum_{n \in \mathbb{Z}} \exp(-n^2 \bar{t}) \leq 2\sqrt{2\pi} \max\left\{1, \bar{t}^{-1/2}\right\}.$$

This implies $\|T\|_{\mathcal{X} \to \mathcal{A}^\lambda(\mathbb{R}/2\pi\mathbb{Z})} \leq 2 \max\left\{1, \bar{t}^{-1/2}\right\}$. The assertions will now follow from Remark 2.1. The conjugate function of λ is $\lambda^*(s) = s^2 \bar{t}$. By Theorem 3.2 we find that (8) holds true for all $0 < \delta \leq \delta_0$, $\delta_0 > 0$ sufficiently small with

$$\gamma(\delta) = \frac{4}{\exp(3\bar{t}) - 1} \exp\left(-\left\lfloor \frac{\pi}{\delta} - \frac{1}{2} \right\rfloor^2 \bar{t}\right)$$

since $m_0 = 1$ and $a = \lambda^*(m_0 + 1) - \lambda^*(m_0) = 3\bar{t}$ in the proof of Theorem 3.2. This yields the claim.

Part 2: One readily calculates that

$$\varphi_p(T^*T)(w)(x) \sim \sum_{n \in \mathbb{Z}} \frac{1}{(\bar{t}n^2)^p} \hat{w}(n) \exp(\mathrm{i}nx),$$

i.e. $\|\varphi_p(T^*T)w\|_{H^{2p}(-\pi,\pi)} \sim \|w\|_{L^2(-\pi,\pi)}$. Thus $f^\dagger \in H^p(-\pi,\pi)$ is equivalent to $f^\dagger = \varphi_{p/2}(T^*T)w$ with $w \in L^2(-\pi,\pi)$. Now Proposition 2.2 with $\mathbb{M} = (-\pi,\pi)$ yields the claim. $\qquad\square$

Theorem 4.2. *Let $\bar{t} > 0$ and T be as in (38) and suppose $f^\dagger \in H^p(-\pi,\pi)$ with some $p > 0$. Suppose furthermore that $\alpha = \alpha(\eta,\epsilon)$ is chosen such that (11) holds true in the limit $\eta,\epsilon,\alpha \searrow 0$.*

Then there exists $\eta_0 > 0$ and a constant $C = C(\bar{t}, T, f^\dagger)$ such that for any noise function $\xi \in L^1(-\pi,\pi)$ fulfilling (2) with $0 < \eta \leq \eta_0$, $\epsilon \geq 0$, the following estimates are valid for Tikhonov regularization (1) with $\mathcal{R}(f) = \|f\|^2_{L^2(-\pi,\pi)}$, $r = 1$:

$$\|\widehat{f}_\alpha - f^\dagger\|^2_{L^2} \leq C\left(\frac{\epsilon}{\alpha} + \frac{\eta^2}{\alpha^2}\exp\left(-\frac{\pi^2}{8\eta^2}\bar{t}\right) + (-\ln(\alpha))^{-p}\right)(1 + o(1)), \tag{40}$$

$$\|T(\widehat{f}_\alpha) - g^\dagger\|_{L^1} \leq C\left(\epsilon + \frac{\eta^2}{\alpha}\exp\left(-\frac{\pi^2}{8\eta^2}\bar{t}\right) + \alpha(-\ln(\alpha))^{-p}\right)(1 + o(1)), \tag{41}$$

as $\eta,\epsilon,\alpha \searrow 0$. The parameter $\alpha = \bar{\alpha}(\epsilon,\eta)$ can be chosen such that

$$\|\hat{f}_{\bar\alpha} - f^\dagger\|^2_{L^2} = \mathcal{O}\left(\max\left\{(-\ln(\epsilon))^{-p}, \eta^{2p}\right\}\right),$$

$$\|T(\hat{f}_{\bar\alpha}) - g^\dagger\|_{L^1} = \mathcal{O}\left(\max\left\{\epsilon, \exp\left(-\frac{\pi^2}{16\eta^2}\bar{t}\right)\eta^{2p+1}\right\}\right)$$

as $\epsilon,\eta \searrow 0$.

Proof. As (11) holds true, we know from Proposition 2.2 that (4) can be applied with $f = \widehat{f}_\alpha$. From Lemma 4.1 we know that we may apply Theorem 2.4 with $\varphi(\tau) = \beta'\varphi_p(\tau)$ and

$$\gamma(\delta) \leq \frac{8\max\{1, \bar{t}^{-1/2}\}}{\exp(3\bar{t}) - 1}\exp\left(-\left\lfloor\frac{\pi}{\delta} - \frac{1}{2}\right\rfloor^2\bar{t}\right)$$

$$\leq 8\frac{\max\{1, \bar{t}^{-1/2}\}}{\exp(3\bar{t}) - 1}\exp\left(-\frac{\pi^2}{\delta^2}\bar{t}\right)(1 + o(1)), \qquad \delta \searrow 0.$$

The function ψ in (15) is given by

$$\psi(\alpha) = \beta'\left(\ln\left(\frac{1}{\beta'\alpha p}\right)\right)^{-p}(1 + o(1)) \leq (\beta')^2 p(-\ln(\alpha))^{-p}(1 + o(1)).$$

Inserting these results into (16) and (17) with $q = q' = 2$ and $b = 2$ yields (40) and (41).

An optimal α can be chosen now depending on which of the first two terms in (40) dominates. If the first term is larger than the second, we choose α_1 to balance the first and the last term in (40), which yields

$$\frac{\epsilon}{\alpha_1} = (-\ln(\epsilon))^{-p}(1 + o(1)), \qquad \alpha_1, \epsilon \to 0.$$

On the other hand, if the second term is larger than the first, we choose α_2 such that the second and third term equal, which gives

$$(-\ln(\alpha_2))^{-p} = \bar{t}^{-p}\left(\frac{4\eta}{\pi}\right)^{2p}(1 + o(1)), \qquad \alpha_2, \eta \to 0.$$

Taking the maximum of both cases yields the proposed bounds. As mentioned in Remark 2.6, the condition (11) is satisfied. $\qquad\square$

15

4.2 An inverse problem in satellite gradiometry

Let us assume, the earth is described by the unit ball $B := \{x \in \mathbb{R}^3 : |x| \leq 1\}$. In geophysics the value of the gravitational potential u on the surface of the earth is of interest as it contains information about the interior of the earth. If the value f of u on ∂B is given, u is the solution to the exterior boundary value problem

$$
\begin{cases}
\Delta u = 0 & \text{in } \mathbb{R}^3 \setminus B, \\
u = f & \text{on } \partial B = SS^2, \\
|u(x)| = \mathcal{O}(|x|^{-1}) & \text{as } |x| \to \infty.
\end{cases}
$$

Using satellites it is possible to measure the second derivative of u in radial direction $r = |x|$ at some distance $R > 1$ from the earth, i.e. the rate of change of the gravitational force:

$$
g = \frac{\partial^2 u}{\partial r^2} \qquad \text{on } RSS^2
$$

The inverse problem of gradiometry then consists in estimating f given g (see [11, Sec. 8.2] and references therein). Representing u by the Poisson formula for the exterior of B, it can be seen (see e.g. [9]) that the forward operator is given by

$$
\begin{aligned}
(Tf)(x) &= \frac{1}{|SS^2|} \int_{SS^2} \frac{\partial^2}{\partial R^2} \left\{ R^{-1} \frac{1 - R^{-2}}{|R^{-1}x - y|^3} \right\} f(y) \, \mathrm{d}y \\
&= \sum_{m=0}^{\infty} \frac{(m+1)(m+2)}{R^{m+3}} (Q_m f)(x) \quad x \in SS^2
\end{aligned}
\tag{42}
$$

with the orthogonal projection Q_m onto the spherical harmonics of order m introduced in (33). This representation of T shows again that the problem to recover f from g is exponentially ill-posed.

Lemma 4.3. *Let T as in (42) with $R > 1$, $\mathcal{X} = L^2(SS^2)$ and $\mathcal{R}(f) = \|f\|_{L^2(SS^2)}^2$.*

1. *T is a bounded mapping from $L^2(SS^2)$ to $\mathcal{A}^\lambda(SS^2)$ for*

$$
\lambda(r) := \begin{cases}
-4 \ln(R - \exp(r)), & 0 \leq r < \ln(R), \\
\infty & \text{else.}
\end{cases}
$$

2. *Assumption 2 holds true with*

$$
\gamma(\delta) = C \delta^{-5/2} R^{-\sqrt{4\pi/\delta} - 4}
\tag{43}
$$

   with C independent of R and δ.

3. *If $f^\dagger \in H^p(SS^2)$, then Assumption 1 holds true with any bounded $D(F)$, $\mathbb{M} = SS^2$, $r = 1$, some $\beta, \beta' > 0$ and*

$$
\varphi(\tau) = \beta' \varphi_{2p}(\tau), \qquad \tau > 0.
$$

Proof. Part 1: First note that for $g_m \in \mathcal{H}_m$ and $n \in \{0, 1, \ldots, \}$ we have

$$
\begin{aligned}
\int_{-1}^{1} p_n(t)(M g_m)(x, t) \, \mathrm{d}t &= \int_{-1}^{1} \frac{p_n(t)}{2\pi\sqrt{1 - t^2}} \int_{\langle x, y \rangle = t} g_m(y) \, \mathrm{d}y \, \mathrm{d}t \\
&= \int_{SS^2} p_n(\langle x, y \rangle) g_m(y) \, \mathrm{d}y \\
&= \frac{4\pi}{2m + 1} (Q_n g_m)(x) \\
&= \delta_{n,m} \frac{4\pi}{2m + 1} g_m(x).
\end{aligned}
$$

As $\{p_m/\sqrt{2m+1}\}$ is an orthonormal basis of $L^2(-1,1)$, we find that $(Mg_m)(x,t) = \frac{p_m(t)}{4\pi}g_m(x)$. Together with (42) we obtain

$$(MTf)(x,t) = \sum_{m=0}^{\infty} \frac{p_m(t)}{4\pi} \frac{(m+1)(m+2)}{R^{m+3}}(Q_m f)(x).$$

Using the bound $|p_m(z)| \leq \exp(mr)$ for $z \in \partial \mathcal{E}_r$ from the proof of Theorem 3.6 and $\|Q_m f\|_{L^\infty(SS^2)} \leq (2m+1)\|f\|_{L^2(SS^2)}$ we obtain

$$\|Tf\|_{\mathcal{A}^\lambda(SS^2)} \leq \sup_{r \geq 0} \sup_{z \in \partial \mathcal{E}_r} \sup_{x \in SS^2} \exp(-\lambda(r))|(MTf)(x,z)|$$

$$\leq \sup_{0 \leq r < \ln(R)} (R - e^r)^4 \sum_{m=0}^{\infty} \frac{p_m(z)}{4\pi} \frac{(m+1)(m+2)}{R^{m+3}}\|Q_m f\|_{L^\infty(SS^2)}$$

$$\leq \sup_{0 \leq r < \ln(R)} \frac{(R - e^r)^4}{4\pi} \sum_{m=0}^{\infty} \frac{(m+1)(m+2)(2m+1)e^{rm}}{R^{m+3}}\|f\|_{L^2(SS^2)}$$

$$\leq CR\|f\|_{L^2}.$$

The last inequality follows by evaluating derivatives of the geometric series, and C is a generic constant independent of R.

Part 2: The supremum of $r \mapsto sr - \lambda(r)$ is attained if $e^r = sR/(4+s)$, so

$$\exp\left(-\lambda^*(s)\right) = \left(\frac{4+s}{4R}\right)^4 \left(\frac{4+s}{sR}\right)^s \leq Cs^4 R^{-s-4}.$$

Now (43) follows from Theorem 3.6.

Part 3: The proof is similar to the heat equation case using [11, Prop. 16]. There it has been shown that

$$f^\dagger = \varphi_p\left(T^*T\right)\omega, \omega \in L^2\left(SS^2\right) \qquad \Leftrightarrow \qquad f^\dagger \in H^p\left(SS^2\right).$$

Thus the spectral source condition is satisfied by assumption, and it follows from Corollary 2.3 that (4) holds true with $r = 1$, and $\varphi = \beta' \varphi_{2p}$ for any bounded $D(T)$. $\qquad\square$

Theorem 4.4. *Let $R > 1$, T as in (42) and $f^\dagger \in H^p\left(SS^2\right)$ with some $p > 0$. Suppose furthermore that $\alpha = \alpha(\eta, \epsilon)$ is chosen such that (11) holds true in the limit $\eta, \epsilon, \alpha \searrow 0$.*

Then there exists $\eta_0 > 0$ and a constant $C = C\left(R, T, f^\dagger\right)$ such that for any noise function $\xi \in L^1(SS^2)$ fulfilling (2) with $0 < \eta \leq \eta_0$, $\epsilon \geq 0$, the following estimates are valid for Tikhonov regularization (1) with $\mathcal{R}(f) = \|f\|_{L^2(SS^2)}^2$, $r = 1$:

$$\|\hat{f}_\alpha - f^\dagger\|_{L^2(SS^2)}^2 \leq C\left(\frac{\epsilon}{\alpha} + \frac{1}{\alpha^2}\frac{1}{R^6}\eta^{-3}R^{-2\sqrt{\pi/\eta}} + (-\ln\alpha)^{-2p}\right)(1 + o(1)),$$

$$\|T(\hat{f}_\alpha) - T(f^\dagger)\|_{L^1(SS^2)} \leq C\left(\epsilon + \frac{1}{\alpha}\frac{1}{R^6}\eta^{-3}R^{-2\sqrt{\pi/\eta}} + \alpha(-\ln\alpha)^{-2p}\right)(1 + o(1)),$$

as $\eta, \epsilon, \alpha \searrow 0$. The parameter $\alpha = \bar{\alpha}(\epsilon, \eta)$ can be chosen such that

$$\|\hat{f}_{\bar{\alpha}} - f^\dagger\|_{L^2(SS^2)}^2 = \mathcal{O}\left(\max\left\{(-\ln\epsilon)^{-2p}, \eta^p\right\}\right)$$

$$\|T(\hat{f}_{\bar{\alpha}}) - T(f^\dagger)\|_{L^1(SS^2)} = \mathcal{O}\left(\max\left\{\epsilon, \eta^{\frac{2p-3}{2}}R^{-\sqrt{\pi/\eta}}\right\}\right)$$

as $\epsilon, \eta \searrow 0$.

Proof. As (11) holds true, we know from Proposition 2.2 that (4) can be applied with $f = \hat{f}_\alpha$. From Lemma 4.3 we know that we may apply Theorem 2.4 with $\varphi(\tau) = \beta'\varphi_{2p}(c\tau)$ and γ as in (43). This yields the claimed error estimates.

To obtain the asserted convergence rates, we may again distinguish if the ϵ term or the η term dominates. Then we balance the dominating term with the pure α term to obtain a choice of α satisfying (11). The maximum of both cases yields the claim. $\qquad\square$

5 Conclusions

We have extended a recent approach for inverse problems with impulsive noise to the case of infinitely smoothing forward operators using standard Sobolev smoothness assumptions for the solution. Remarkably, one obtains high order polynomial rates of convergence in the size η of the corrupted domain, even though the underlying problem is exponentially ill-posed in the classical sense. We examined two exponentially ill-posed problems arising in PDEs and showed that our analysis can be applied.

Our study gives rise to several further interesting questions: One concerns the extension of our analysis to more general domains in $\mathbb{R}^d$. A possible strategy could be the use of local averages as introduced here for the case of SS^2. However, boundaries would cause technical difficulties, and the choice of approximating subspaces is not obvious. So far we have only studied a priori parameter choice rules, i.e. the solution smoothness characterized by φ has to be known for choosing α. The popular discrepancy principle does not apply in our context (as only $\|\xi\|_{L^1(\mathbb{M}\setminus\mathbb{P})}$ is small, but $\mathbb{P}$ is unknown), and thus other a posteriori parameter choice rules would be of interest for practical applications. Other open questions include the optimality of the error bounds and an analysis of iterative regularization methods for nonlinear inverse problems with impulsive noise.

Acknowledgement

References

[1] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Probl.*, 20(5):1411–1422, 2004.

[2] R. H. Chan, C. Hu, and M. Nikolova. An iterative procedure for removing random-valued impulse noise. *IEEE Signal Proc. Let.*, 11(12):921–924, 2004.

[3] C. Clason and B. Jin. A semismooth Newton method for nonlinear parameter identification problems with impulsive noise. *SIAM J. Imaging Sci.*, 5(2):505–536, 2012.

[4] C. Clason, B. Jin, and K. Kunisch. A duality-based splitting method for ℓ^1-TV image restoration with automatic regularization parameter choice. *SIAM J. Sci. Comput.*, 32(3):1484–1505, 2010.

[5] C. Clason, B. Jin, and K. Kunisch. A semismooth Newton method for L^1 data fitting with automatic choice of regularization parameters and noise calibration. *SIAM J. Imaging Sci.*, 3(2):199–231, 2010.

[6] J. Flemming. *Generalized Tikhonov regularization and modern convergence rate theory in Banach spaces*. Shaker Verlag, Aachen, 2012.

[7] J. Flemming and B. Hofmann. A new approach to source conditions in regularization with general residual term. *Numer. Funct. Anal. Optimiz.*, 31:254–284, 2010.

[8] J. Flemming, B. Hofmann, and P. Mathé. Sharp converse results for the regularization error using distance functions. *Inverse Probl.*, 27(2):025006, 2011.

[9] W. Freeden, F. Schneider, and M. Schreiner. Gradiometry—an inverse problem in modern satellite geodesy. In *Inverse problems in geophysical applications (Yosemite, CA, 1995)*, pages 179–239. SIAM, Philadelphia, PA, 1997.

[10] M. Grasmair. Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Probl.*, 26(11):115014, 2010.

[11] T. Hohage. Regularization of exponentially ill-posed problems. *Numer. Funct. Anal. Optimiz.*, 21:439–464, 2000.

[12] T. Hohage and F. Weidling. Verification of a variational source condition for acoustic inverse medium scattering problems. *Inverse Probl.*, 31:075006, 2015.

[13] T. Hohage and F. Werner. Convergence rates for inverse problems with impulsive noise. *SIAM J. Numer. Anal.*, 52(3):1203–1221, 2014.

[14] B. Jin. A variational Bayesian method to inverse problems with impulsive noise. *J. Comput. Phys.*, 231(2):423–435, 2012.

[15] Q. Jin and L. Stals. Nonstationary iterated Tikhonov regularization for ill-posed problems in Banach spaces. *Inverse Probl.*, 28(10):104011, 2012.

[16] T. Kärkkäinen, K. Kunisch, and K. Majava. Denoising of smooth images using L^1-fitting. *Computing*, 74(4):353–376, 2005.

[17] R. Kress. *Linear integral equations*. Springer-Verlag, New York, 1999.

[18] Y.-R. Li, L. Shen, D.-Q. Dai, and B. W. Suter. Framelet algorithms for de-blurring images corrupted by impulse plus Gaussian noise. *IEEE Trans. Image Process.*, 20(7):1822–1837, 2011.

[19] C. Müller. *Spherical harmonics*. Springer, 1966.

[20] M. Nikolova. Minimizers of cost-functions involving nonsmooth data-fidelity terms. Application to the processing of outliers. *SIAM J. Numer. Anal.*, 40(3):965–994 (electronic), 2002.

[21] M. Nikolova. A variational approach to remove outliers and impulse noise. *J. Math. Imaging Vision*, 20(1-2):99–120, 2004. Special issue on mathematics and image analysis.

[22] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Applied Mathematical Sciences. Springer, 2008.

[23] E. M. Stein and R. Shakarchi. *Fourier analysis: An Introduction*. Princeton University Press, 2003.

[24] E. M. Stein and G. L. Weiss. *Introduction to Fourier analysis on Euclidean spaces*, volume 1. Princeton university press, 1971.

[25] G. Szegö. *Orthogonal polynomials*. AMS, 1959.

[26] F. Werner and T. Hohage. Convergence rates in expectation for Tikhonov-type regularization of Inverse Problems with Poisson data. *Inverse Probl.*, 28(10):104004, 2012.

[27] J. Yang, Y. Zhang, and W. Yin. An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise. *SIAM J. Sci. Comput.*, 31(4):2842–2865, 2009.

[28] W. Yin, D. Goldfarb, and S. Osher. The total variation regularized L^1 model for multiscale decomposition. *Multiscale Model. Simul.*, 6(1):190–211 (electronic), 2007.